基于影响度及状态预测的多 agent 协作算法 *

郑延斌，马光富，王林林，席鹏雪

(河南师范大学 计算机与信息工程学院，河南 新乡 453007)

摘　要：针对灾难环境中多 agent 协作问题进行了研究，提出了一种基于影响度与状态预测的多 Agent 协作算法。首先该算法根据协作任务对信息的需求，使用影响度函数对 agent 感知到的信息进行处理；其次利用预测算法对任务的后续状态和 agent 的行为进行预测并根据预测结果制定协作策略；最后执行协作任务的 agent 根据动作效果和触发条件动态调整协作策略。为了验证算法的有效性，在 Unity3D 中搭建仿真平台，对比不同协作算法的收敛率、救援人数和整体得分，结果表明该算法的收敛速度快、救援人数多和整体得分最优，可以效地指导 agent 间的协作，能给实际救援协作策略的制定提供理论支持。

关键词：多 agent；信息处理；状态预测；协作策略；动作效果

中图分类号：TP301.6　　doi: 10.3969/j.issn.1001-3695.2018.03.0208

Multi-agent cooperation rescue algorithm based on influence degree and state prediction

Zheng Yanbin, Ma Guangfu, Wang Linlin, Xi Pengxue

(*College of Computer & Information Engineering Henan Normal University*, *Xinxiang Henan* 453007, *China*)

Abstract: Aiming at the problem of multi-agent cooperation in disaster environment, a multi-agent cooperation algorithm based on influence and state prediction is proposed. Firstly, the algorithm uses the influence degree function to process the information perceived by the Agent based on the information requirements of the collaborative task. Secondly, the predictive algorithm is used to predict the subsequent state of the task and the behavior of the agent, and a collaborative strategy is formulated based on the prediction result. Finally, the Agent performing the collaborative task dynamically adjusts the collaboration strategy according to the action effect and the trigger condition. In order to verify the effectiveness of the algorithm, a simulation platform was built in Unity3D to compare the convergence rate, the number of rescuers, and the overall score of different collaborative algorithms. The results show that the algorithm has the faster convergence speed, the better number of rescuers and the higher of overall score, which can effectively guide the collaboration between Agents and provide theoretical support for the development of practical rescue coordination strategies.

Key words: multi-agent; information filtering; state prediction; collaborative strategy; action effect

0　引言

多 agent 系统（multi-agent system, MAS）是分布式人工智能领域研究的热点，目前动态环境下多 agent 协作是 MAS 研究的核心问题之一[1~3]。如文献[4]提出了一种适用于海港防护的多 agent 分散部署策略，在信息噪声大和通信间断的情况下实现了 agent 的调度、部署、拦截导弹攻击等任务，但此算法对快速变化信息的处理效率较低。文献[5]使用蒙特卡洛树搜索算法实现无人机直接的协作，该算法通过计算邻居结点的值构建关系图并根据奖励函数制定联合动作，当结点关系发生突变时将降低算法的性能可能导致协作任务无法完成。文献[6]提出了一种预测系统并应用到多 agent 的协作中，该系统将遗传算法和进化论相结合，通过计算环境的变化对 agent 的进化过程进行预测，当环境出现突变时此算法无法完成相关预测。文献[7]通过对个体的意图、行为动作等进行识别和预测进而得到群体的联合意图、行为动作等，并为协作策略的制定提供有利的支持，但其忽视了环境变化对个体行为动作的影响。国内学者如文献[8]对 MA-PDDL 进行了改进，在语言中加入了连续规划元素，提出了一种基于连续规划的协作算法，该算法避免了协作过程中因环境改变而导致任务失败的情况，但是 agent 在执行动作前都要更新系统环境，其时间花费过大。文献[9,10]通过融合系统对环境信息进行处理与合成得到目标的综合状态，agent

收稿日期：2018-03-28；修回日期：2018-05-07　　基金项目：河南省基础前沿资助项目（142300410349）；河南省软科学资助项目（14240041001）

作者简介：郑延斌（1964-），男，河南内乡人，教授，博士，主要研究方向为虚拟现实、多智能体系统、博弈论（zybcgf@163.com）；马光富（1990-），男，硕士研究生，主要研究方向为虚拟现实、多智能体系统；王林林（1993-），女，硕士研究生，主要研究方向为虚拟现实、多智能体系统；席鹏雪（1993-），女，硕士研究生，主要研究方向为虚拟现实、多智能体系统.

根据状态信息制定协作策略，但在动态环境中其融合过程信息运算量巨大且复杂、寻找全局最优解困难。文献[11]使用粒子滤波算法对带有噪声的信息进行处理并得到目标状态，但其算法复杂度较高直接影响策略的制定。文献[12]提出了一种基于状态预测的多智能体动态协作算法，该算法通过给信息添加紧急度对协作信息进行处理并对其后续状态进行预测，此方法能减少信息的处理量且筛选效率较高，但个体在接收紧急度高的信息的同时可能会丢弃与协作任务的密切相关的重要度较低的信息，影响协作策略的正确性。

灾难环境复杂多变且具有突发性强、破坏力大、易产生次生灾害、影响深远、救援难度大等特点，这使得上面研究者提出的算法不能很好解决灾后多 agent 协作救援的问题，因此本文提出了一种基于影响度及状态预测相结合的多 agent 协作算法，该算法首先根据协作任务对信息的要求，使用影响度函数有针对性的对感知到的信息进行处理，得到与任务密切相关的信息忽略无关信息；其次引入状态预测算法，对协作任务的发展趋势和个体的行为进行预测，并根据预测结果制定协作策略；最后协作 agent 根据协作策略和任务的具体要求选取对应动作执行任务，agent 根据动作效果对协作策略进行修改，实现协作过程的动态调整，保证了系统的协作效率和正确率。仿真实验表明该算法能够综合环境、预测、规划等信息完成灾后协作救援任务，为实际协作决策的制定提供了理论和技术支持。

1　信息筛选

首先灾难环境通常包括大量倒塌和受损的建筑物、火灾、伤员、阻塞的道路等动态变化的信息,且具有信息种类多、噪声大、离散程度高、处理方式不一致等特性，此外这些信息与协作任务的关系不同产生的影响也不同，因此需要从复杂信息中筛选出与救援人密切相关的信息；其次协作 agent 感知和接收信息的能力有限，所以需要从含有噪声的信息中提取出实时有效的信息忽略无关信息，减少个体间协作信息的传输量，提高协作的效率。因此本文使用影响度函数对协作信息进行筛选。

1.1　影响度函数

影响度函数分为两层如图 1 所示，式（1）~（3）作为影响度函数的第一层，其作用是确定救援任务的类型信息，如灾害的种类、影响面积、受灾人数等；第二层是补充信息层，它由许多灾害类构成且每个类中包含大量的内容不同的信息筛选函数，其作用是根据式（1）中的 dis 调用与之对应的灾害类补充与任务密切相关的信息。

$$D = dis * \alpha \quad (1)$$

$$S = \pi * r^2 \quad (2)$$

$$N = P * F \times \Upsilon \quad (3)$$

其中：dis 表示灾害类型包括地震、泥石流、洪水、火灾等自然灾害或突发公共事件。α 表示灾害对应的等级，取值为常数，其值越大代表灾害带来的损害越大。S 表示灾害的影响面积，r 为灾源 (x_0,y_0) 到离灾源最远的受灾点 (x_i,y_i) 的距离。N 表示个体或建筑物受到的损害值。当 $p=p_1$ 时表示受灾人数，当 $p=p_2$ 时表示受损建筑物数，F 代表其受到损害的初值，Υ 表示点 $r(x_i,y_i)$ 的人员或建筑物等受到灾源影响的系数，值越接近 1 表示其受到的危害度越大。f 是调节因子。

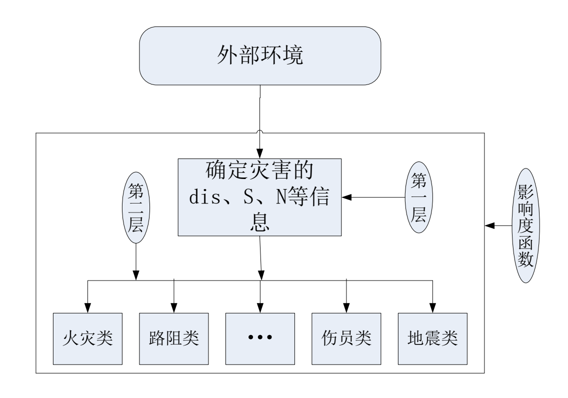


图 1　影响度函数结构

由于不同类型的灾难其构成特点和危害不同，此时协作救援所需要的信息也不同。因此需要使用第二层对任务的详细信息进行补充，下面以火灾类和伤员类以及路阻类为主介绍信息筛选流程。当 Agent 执行救火任务时根据式（4）~（6）获取火灾的具体信息。

$$f(area) = R_0 \times Mat \times M_0 \times K_\varphi \times K_w \times s \quad (4)$$

$$C(x,y,z,t,H) = \frac{2Q}{(2\pi)^{3/2}\sigma_x\sigma_y\sigma_z}\exp[-\frac{(x-V_t)^2}{2\sigma_x^{\ 2}}]\times$$
$$\exp\frac{y^2}{2\sigma_y^2}\{\exp[-\frac{(z-H)^2}{2\sigma_z^{\ 2}}]+[-\frac{(z+H)^2}{2\sigma_z^{\ 2}}]\} \quad (5)$$

$$TL = \int_1^2 C_i^{\,n}(t)dt \quad (6)$$

根据式（4）获取如下具体的信息，火焰的蔓延速度 R_0，燃料的引燃程度 Mat，燃料的初始质量 M_0，火灾点的温度、湿度、建筑物结构、灾员分布等信息 K_φ，通风因子 K_w。式（5）可得到有毒气体的信息如空气中有毒气体浓度 C[13]，毒气高度 H，火灾源产生毒气的量 Q，传播时间 t，x、y、z 方向的扩散系数 $\sigma_x\sigma_y\sigma_z$。式（6）可以得到伤员受毒气危害的信息，如毒气的毒负荷 TL，伤员周围毒气浓度 C_i，接触时间 t 等信息，其中 n 为幂指数且对不同的毒气取值不同。

发现伤员时通过使用式（7）（8）可以获取与伤员密切相关的信息。

$$H_{t+1} = H_0 - d_0 - r(t) \quad (7)$$

$$r(t+1) = \sum r^{bt}(t) + k^{dis} \times r(t)^2 + l^{dis} + n \quad (8)$$

式（7）可以获取伤员生命值的具体信息，s 如个体的初始健康值 H_0，灾害对个体的初始损伤值 d_0，个体健康值的下降速率 $r(t)$；式（8）可以获得 $r(t)$ 的具体信息，如个体的受伤类型 bt，k^{dis} 和 l^{dis} 表示在不同灾害类型中的比例因子，n 为高斯噪声。

接受协作任务的个体调用路阻类，使用式（9）~(11)获取通往目标点道路上的障碍物信息以及到达目标点的预期时间。

$$T_{mi}=\frac{N_b_i}{v_{m_b_i}} \tag{9}$$

$$T_i=S_i/V \tag{10}$$

通过式（9）可以获取障碍物的类型 b_i、数量 N_b_i、清障速度 $v_{m_b_i}$、清障时间 T_{mi} 等具体信息。式（10）得到清障车到达障碍物的预期时间，其中 S_i 表示清障车辆到达障碍物的最优路径的长度[14],V 是清障车的平均移动速度。

$$T=\begin{cases} t_i, & t_i>T_i\text{且}\ (t_i-T_i)>N_b_i/v_{m_b_i} \\ t_i+\dfrac{N_b_i-(t_i-T_i)\times v_{m_b_i}}{v_{m_b_i}}, & t_i>T_i\text{且}\ (t_i-T_i)<N_b_i/v_{m_b_i} \\ T_i+\dfrac{N_b_i}{v_{m_b_i}}, & t_i<T_i \end{cases} \tag{11}$$

$$T_c=\sum_{j=1}^{j=m}T \tag{12}$$

规定场景中的救援车辆以 V 匀速运动，通过比较 Agent 到达障碍物 b_i 的时间 t_i 和清障车到达障碍物 b_i 的时间 T_i 得到通过障碍物所用的时间。如式（11）所示当 $t_i>T_i$ 且 $(t_i-T_i)>N_b_i/v_{m_b_i}$ 时其通过障碍物的时间为 t_i，当 $t_i>T_i$ 且 $(t_i-T_i)<N_b_i/v_{m_b_i}$ 时其通过障碍物的时间为 $t_i+\frac{N_b_i-(t_i-T_i)\times v_{m_b_i}}{v_{m_b_i}}$；当 $t_i<T_i$ 时其通过障碍物的时间为 $T_i+\frac{N_b_i}{v_{m_b_i}}$。使用式（12）可以得到救援车辆到达目标点的时间 T_c 其中 j 为经过的障碍物数目。

使用影响度函数可以获取与协作任务密切相关的信息,不对这些信息进行处理直接使用，会造成场景中同时存在大量相似的信息，给接收信息的个体造成困扰，如果接收错误的信息将导致协作任务失败。因此，需要对筛选出的信息进行加工使其与协作任务唯一对应，避免个体接收重复的或与协作任务无关的信息。

1.2 信息加工

通过使用影响度函数可以筛选出与协作任务密切的相关的信息，但场景中可能存在多个类似的任务，这就造成场景中同时存在一些相似的信息，如果不对这些信息进行加工处理将会对个体的协作造成影响。如系统中同一时刻可能存在多个伤员信息，如果协作个体无法辨别这些信息或接收错误的信息将导致协作任务无法进行；当其选择全部接收信息时一方面增加了个体的通信压力和花费大量的时间，另一方面这些信息对应的预测结果不同，导致个体无法给出正确的收益值丧失动态修改策略的能力。针对这种情况本文提出采用任务编号加协作信息的方式对信息进行加工，使其变成与协作任务唯一对应的信息。个体在接收信息时只需选择与自身任务编号相同的信息即可。

任务编号由时间、地点和随机数三部分组成，编号中的时间是指搜救智能体第一时刻发现任务所对应的系统时间；地点是指协作任务在场景中的位置；随机数是由系统随机产生的其作用将同时同地的不同任务进行区分。传输信息时发送者需要将协作信息和它对应的任务编号一同发送，接受任务的个体只需接收和自身任务编号相同的信息并将其保存。

1.3 筛选算法

影响度函数筛选信息的主要算法是基于信息对当前任务的影响度，对场景中的信息进行提取得到与救援任务密切相关的信息，忽略无效的信息。信息筛选的主要算法如下：

算法开始

a)在场景中搜索，发现任务，通过传感器获取信息，执行 b)；

b)调用第一层中的相关公式如（1）得到灾害类型 *dis* 和受灾等级 *α*，调用式（2）得到影响面积 *S*，调用式（3）得到人员或房屋受到的伤害值 *N*；当 *α*>*λ* 或满足 *α*<*λ* 且 *N*>0 时则执行 c)；否则返回 a)；

c)调用第二层中与 *dis* 对应的灾害类得到任务的具体信息，如调用式（4）~（6）等获取火灾的具体信息，调用式（7）-（8）等得到伤员的具体信息，执行 d）；

d)调用式（9）~（12）等得到救援路径上路阻的具体信息，执行 e)；

e)将筛选出的信息进行加工，并保存；

算法结束

在复杂动态的 MAS 中，使用影响度函数对协作信息进行筛选，在提高信息正确度的同时减少了个体间发送和接收的信息量降低系统的通信压力。

2 状态预测

为了保证协作策略和个体动作的正确性，算法必须能够对任务的发展过程进行及时准确的预测。当前对个体行为和动作的预测技术已十分成熟，但缺少有效的方法对任务和环境的发展与变化趋势进行预测。为此提出一种新的预测算法，首先使用缓存中的数据对任务的发展趋势、个体的行为、环境的变化等进行初步预测，并根据预测结果制定协作策略。其次计算动作效果，得到其对任务、环境、其他个体等的影响，根据影响大小调整协作策略。

缓存中的信息 X_{tn} 和系统当前状态 U_{tm} 作为状态预测函数的信息输入，对协作任务的后续状态、环境变化、个体行为等进行预测，得到 $t+T_c$ 时刻的相关状态 U_{t+T_c}。状态预测主要函数公式如下：

$$U_{t+T_c}=f(X_{tn},U_{tm})\pm(\delta\times T_c) \tag{13}$$

$$p^{'}=f(x_{t-1},u_{t-1}) \tag{14}$$

$$p=f(x_{t-1},u_{t-1}) \tag{15}$$

$$\overline{E}=\frac{\sum_{i=t-k}^{t}\left|p^{'}-p\right|}{k} \tag{16}$$

$$\delta=\sqrt{\frac{\sum_{0}^{N}(p^{'}-p)^2}{N}} \tag{17}$$

其中：p' 为系统的预测值，p 为 $t+T_c$ 时刻系统的实测值，$|p'-p|$ 表示误差值，$\overline{E}$ 表示前 k 个时刻误差的平均值，δ 为剔除 N-k 个较大误差值后的标准误差。预测函数 $f()$ 针对不同的事件或个体其定义不同。状态预测的主要算法如下：

算法开始：

1.输入缓存中的信息 X_{tn} 和 U_{tm}，执行 2；

2.是用式（14）和（15）获取前 *t-k* 个时刻的预测值和实测值，算出其误差值| p' -p|，执行 3；

3.求出前 k 个时刻误差的平均值 $\overline{E}$，执行 4；

4.剔除偏离平均误差较大的预测值，执行 5；

5.求出剩余 N 个误差值的标准差 δ，执行 6；

6.调用式（13）得到状态 U_{t+T_c}；//此状态是一定误差范围内的近似状态

算法结束

剔除误差较大的值主要是防止一些异常时刻的预测值对预测结果的影响。状态预测算法根据上述原则得到个体、任务、环境等的后续状态 U_{t+T_c}。由于环境、个体动作、行为等都会对协作产生影响，因此正确高效的协作策略对 Agent 做出合理的动作，确保协作任务的顺利进行至关重要，是个体间协作不可缺少的一部分。

3 协作策略

要保证 Agent 高效的完成协作任务，需一套完善的协作流程，为救援行动的顺利开展创造条件。目前对多 agent 协作算法的研究较为成熟[15]，本文在现有协作算法基础上加入触发类使算法适用于灾难协作救援，该算法使用触发类对协作策略进行动态调整，避免任务出现死锁使其适用于灾难协作救援。

3.1 策略制定

发现任务的搜索 agent 首先使用信息筛选算法得到与协作任务密切相关的信息并保存，其次使用预测算法得到任务的后续状态。搜索 agent 将筛选出的信息、当前状态、预测状态等传递给总指挥 agent，它根据现有信息确定协作团队、制定初步的协作策略并将其发给团队指挥和协作 agent。接受协作请求的 agent 根据当前信息选取合适的动作，执行后需对动作效果进行评价并判断是否达到触发条件。协作流程图如图 2 所示。

触发类中根据不同的问题设置对应的触发条件，agent 在执行任务的过程中达到触发条件就对协作策略、动作、团队等进行调整。触发条件主要分为三类：第一类主要对 agent 前往任务点的过程中遇到问题如清障时间延长、受到意外损伤、救援车辆抛锚等问题进行规划；第二类是针对协作过程中的问题如任务量变化、环境变化、个体间出现冲突等。第三类主要为解决团队间的资源、时间、空间等冲突进行的规划。

3.2 代价函数

空闲个体根据相关信息计算完成任务的预期代价，并根据 *Cost* 值选择是否接受协作任务，代价函数如式（13）所示。

$$B'_{ij}(T)=\sum_{t=1}^{t=s}B_{ij}(\mathrm{t})*x_{ij} \quad (18)$$

$$x_{ij}=\begin{cases}0,个体\mathrm{i}放弃任务\mathrm{j}\\1,个体\mathrm{i}完成任务\mathrm{j}\end{cases}$$

代价函数是指个体 i 在完成任务 j 的过程中消耗的能力。$B_{ij}(\mathrm{t})$ 为个体在 t 时刻执行选定动作需消耗的能力值，$B'_{ij}(T)$ 为完成协作任务的预期代价值。

$$Cost=A_i-B'_{ij}(T) \quad (19)$$

其中 A_i 为第 i 个 Agent 的现有健康值决定其完成任务的能力，*Cost* 表示个体完协作任务预期剩余健康值。当 *Cost<20* 时表示个体在执行任务时受到严重伤害，放弃协作任务且须补充能力。当时 *Cost>20* 时个体接受任务加入协作团队。

$$Cost'=A_i-[B'_{ij}(T)+s(t)] \quad (20)$$

该式用来对动作代价进行计算，s(t)代表 Agent 执行动作时受到的突发伤害，$Cost'$ 表示个体完成规定动作后剩余的健康值，$B'_{ij}(T)+s(t)$ 表示完成动作的实际代价值。

$$R=\begin{cases}-2,个体故意放弃任务j\\0,个体i受伤退出任务j\\1,个体i完成任务j\end{cases} \quad (21)$$

R 作为奖赏值，系统根据 Agent 完成任务的具体情况返回给 Agent 对应的值，总指挥 Agent 选择当前系统内 R 最大的空闲个体作为团队指挥。

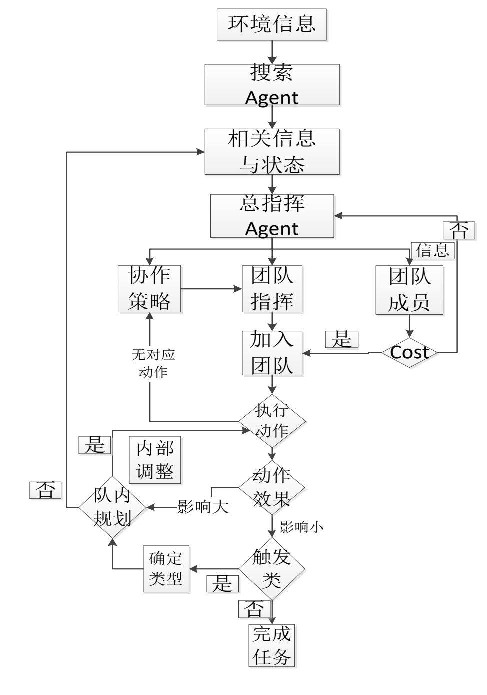


图 2 协作流程

3.3 动作效果

个体完成动作后需对其产生的效果进行评价，并根据评价结果修改协作策略，消解协作过程中产生的冲突。采用三元组描述动作。

定义 1 动作 O 是三元组<A，Q，E>，A 为执行动作的个体，Q 为前提条件集，E 为动作收益。

定义 2 前提条件 Q 是五元组<I，O，S，G，U>，I 为初

始状态即U_t，O 为个体的动作集，S 为动作的相关约束条件，G 为执行选定动作产生的新状态，U 为预测的状态即U_{t+T_c}。

动作效果的评价分为两部分，使用式（20）计算完成动作的实际代价值；主要是比较差值|G -U|与标准误差δ的大小关系，进而得到其对任务、其他个体、环境等产生的影响，并根据影响大小修改策略。对动作效果进行评价的主要算法如下：

算法开始：

a) 根据 I 和 S 以及策略等信息选择动作 O，执行 2；

b) 执行动作 O，调用公式（19）当受到意外伤害且$Cost' < 20$时则放弃当前任务并反馈当前信息修改策略，否则获取当前状态 G 执行 3；

c) 当|G -U|> δ 时执行 4，否则执行 7；

d) 获取具体的影响信息，执行 5；

e) 满足队内调整则执行 6，否则执行 8；

f) 队内调整并规划新动作，若动作集满足要求则执行 8，否则执行 7；

g) 若不满足要求则修改策略更新动作集；

h) 若任务未完成则选取新动作执行 1；否则执行 9；

i) 完成任务，动作评价结束；

算法结束

通过该算法对动作的效果进行评价，在保证动作集正确性的前提下实现了对协作策略的动态调整，避免因策略更新过慢对任务或个体产生影响，提高任务的执行效率。

4　协作救援算法

由于灾难救援任务具有复杂多变的特性且单个 agent 的能力有限，因此需要多个 agent 相互协作共同完成救援任务，最大限度减小灾难带来的损失。此时支持个体间高效协作的救援算法就显得尤为重要，本系统中救援智能体间协作算法如下：

a)搜索 agent 在场景中移动搜索任务；

b)发现任务，通过自身携带的传感器获取任务和环境的相关信息；

c)通过影响度函数第一层得到任务的类型 *dis* 和等级 α 信息；如果不需协作救援则忽略此信息，执行 a)，否则执行 d)；

d)确定受灾面积 *S*、受灾人数、受损房屋等信息执行 e)；

e)使用影响度函数第二层中与 *dis* 对应的灾害类获取任务的具体信息，执行 f)；

f)将筛选出的信息进行加工并存入缓存，执行 g)；

g)状态预测算法利用缓存中的信息对任务、个体、环境等的状态进行预测，得到预测结果并发送给总指挥 agent，执行 h)；

h)总指挥 agent 根据缓存中的信息和预测结果制定协作策略、确定团队成员和指挥 agent，执行 i)；当收到个体修改协作策略的信息时则更新协作策略，执行 k)；

i)向团队个体和指挥 agent 发送协作信息，执行 j)；

j)个体接收协作信息，根据式（19）计算执行相应任务的预期代价值，当 *cost<20* 时放弃当前任务并将结果返回给总指挥 agent，执行 h)；否则加入协作团队，执行 k)；

k)团队指挥 agent 接收协作策略，并将策略发送给接受任务的个体，执行 l)；

l)Agent 按照协作策略选取动作集中的动作并执行，完成动作后使用算法对动作的效果进行评价，当其影响值较小时执行 o)，否则执行 m)；

m)当其满足队内调整时则执行 n)，否则返回 h)；

n)进行队内调整，并更新动作集执行 o)；

o)判断是否达到触发条件，若达到触发条件则执行 m)，否则执行 p)；

p)完成任务 agent 获得奖赏值 R 执行 q)，否则执行 l)；

q)Agent 将现有状态发送给总指挥 Agent，若达到结束条件则执行 r)，否则接受新任务执行 j)；

r)仿真结束，清除场景中和缓存中的信息。

该协作算法首先使用影响度对信息进行处理降低了系统的通信量同时提高了信息的准确性，其次状态预测函数使用缓存中的信息得到系统的后续状态，并根据预测结果制定协作策略指导个体的协作行为，最后利用触发类和动作评价实现协作的动态调整，保证了策略的高效性。

5　实验仿真与分析

5.1　场景建模

本文在 500m*500m 的灾难场景（如图 3）中模拟个体间的协作过程并对所提算法的性能进行验证，场景中模型信息和作用如表 1 所示。

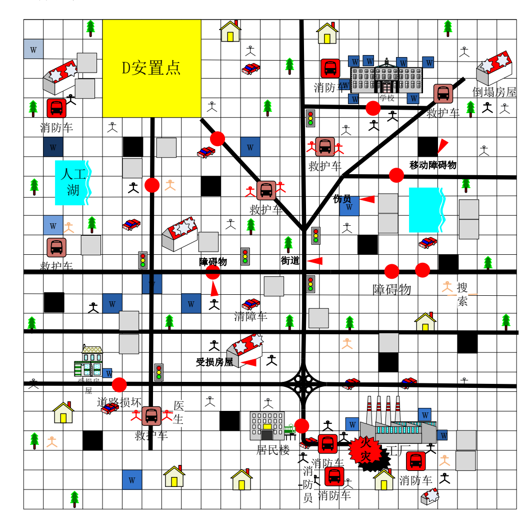


图 3　二维仿真场景

表 1　实验中用到的模型信息

模型名称	数量（个）	相关属性	主要任务及分工
搜索 Agent	20	搜索	处理信息、预测状态
通信 Agent	20	通信	传输协作信息
指挥 Agent	1	指挥	制定策略、确定人员
消防 Agent	150	0.5m/s	灭火、营救伤员
警察 Agent	100	0.5m/s	清障、维持秩序

医生 Agent	80	0.5m/s	救助伤员
救援车	各 30	5m/s	灭火、清障、救护
伤员 Agent	150	0m/s	等待救援
房屋	50	受损	倒塌、掩埋人员
学校	1	倒塌	倒塌、掩埋人员
工厂	1	着火	倒塌、掩埋人员
障碍物	若干	若干	破坏道路、移动
其他	若干	若干	增加真实性

场景主要信息如下：黄色区域 D 为 50m*50m 的安置所，作用是安置伤员和存放物资；黑线条表示 4m 宽的道路，红色圆圈代表静态障碍物且颜色深浅代表不同数量和类型的障碍物，黑色方块代表移动障碍物且无碰撞条件下每个时间步内可随机地向上、下、左、右任意方向移动一个栅格；灾害类型 *dis* 和等级 α 共同决定建筑物的受损情况；蓝色区域表示伤员，且颜色不同表示受伤类型和伤情不同；假设救援车辆在场景中以 5m/s 匀速运行，每辆救护车配备一名医生 Agent，可以根据任务要求调整医生数量；灭火消防车配有 6 名灭火 Agent，同时携带 5000L 水和 1000L 泡沫，其直流射程分别为 65m 和 60m，假设每辆车灭火范围为 100 ㎡且灭火速度为 2 ㎡/s，登高消防车的灭火高度为 100m，升高臂顶端救援平台可一次救援 4 名成年人，工作直径为 20m 且带有消防水炮；清障车主要用来清除道路上的障碍物且配备司机 Agent 和指挥 Agent 各一名，假设其清障速度为 3 m³/s。

5.2　Agent 建模

为了提高仿真的效率，需要明确个体的功能，对不同的 agent 进行建模并对其行为进行约束。实验中的总指挥 agent 的主要功能如下：a）利用相关状态和已有的分配算法对任务进行分配，确定所需 agent 的种类和数量;b）制定协作策略，确定团队指挥 agent 并向协作 agent 发送协作信息;c）收到 agent 反馈的信息时及时更新策略;d）存储系统内 agent、环境、救援车辆等状态信息。团内指挥 agent 主要功能是与总指挥通信，接收并发送协作策略，负责队内规划。

为了提高数据的收集速度和质量，我们将相关传感器布置在搜索 agent 身上，当其发现任务时可以及时感知信息降低了获取信息的时间。搜索 agent 的主要功能如下：a）携带传感器感知环境、其他 agent、任务等信息;b）使用影响度函数筛选信息，并对协作进行加工;c）使用函数预测任务和个体的后续状态;d）接收和传递信息。

消防 agent 主要负责灭火和营救伤员，警察 agent 主要指挥清障车和维持秩序，为其他 agent 提供帮助。医生 agent 主要负责救治伤员，通信 agent 主要负责个体间的信息传输。Agent 有三个行为状态及空闲、忙碌、执行任务，且有 8 个移动方位，移动速度为 0.5 m/s，发生碰撞时健康值减少 20，受到火烧、掩埋、毒气等意外伤害时其健康值根据具体情况减少；且 Agent 在协作时可根据任务需求修改动作集变换其职能。

5.3　仿真结果

动态环境中，agent 在协作前不具备全局动态和静态环境信息，在协作过程中也无法获取动态障碍物的移动信息，相同实验背景下，采用同样的算法参数，进行第一组实验：agent 按照[5]中的算法进行协作； agent 运用文献[6]所提算法协作；agent 使用本文提出的算法进行协作；实验中 agent 的最终目的是一致的，但算法的收敛速度有所不同，实验结果如图 4 所示。

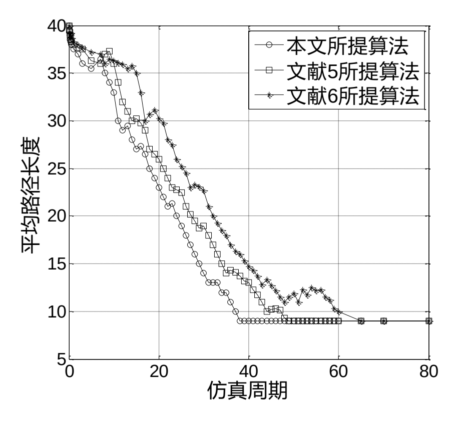


图 4　算法收敛性

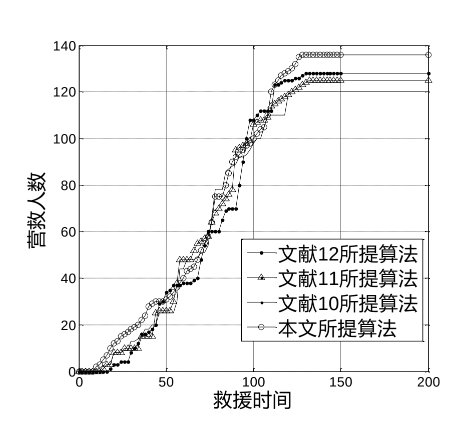


图 5　救援人数

由于单次仿真结果具有随机性，所以本文取 100 次仿真的平均值作为实验结果，仿真结果表明本文算法的收敛速度比其他两种算法都快，这种差异是由 agent 在协作中有无信息处理和状态预测造成的。文献[5]所提算法，agent 在协作救援过程中通过相关计算得到联合动作，但缺少预测算法对相关状态进行有效的预测，当任务出现突变时将影响算法收敛速度和协作效率。文献[6]中的算法，在 agent 协作过程中加入了状态预测，在协作中可以对障碍物、个体、任务的状态进行预测，但其信息处理效率低，导致算法的收敛率和任务的完成率较低。本文考虑上述问题提出将信息处理与状态预测相结合并改进，明显提高了算法的收敛速度。原因首先是本算法没有盲目对所有感知到的信息进行处理，而是只筛选对协作任务有影响的信息，其次利用筛选出的信息预测相关状态并制定协作策略，受到干扰后能够快速找到收敛路径，从而提高协作效率。

在相同信息背景下，进行第二组实验：agent 运用文献[10]中的算法进行协作； agent 采用文献[11]中的算法进行协作；agent 根据文献[12]中的算法进行协作；agent 按照本文所提算

法进行协作；得到四种算法的救援人数如图 5 所示，总体得分如图 6 所示。

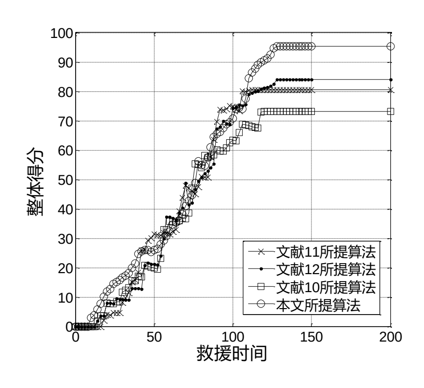


图 6　整体得分

从图 5 的结果可以得出本文算法的救援人数高于文献[12]是因为得到的协作信息更全面，提高了预测结果的正确率，且利用动作评价和触发类实现协作策略的动态调整，保证任务能够顺利完成。文献[11]因信息处理算法的复杂度较高进而降低了协作的效率。而文献[10]难以得到最优的协作策略，导致协作效率低下。图 6 结论进一步说明本文所提算法的性能优于其他算法。

6　结束语

本文提出了一种适用于多智能体协作救援的算法，该算法将影响度函数和状态预测相结合并应用于协作救援个体。首先，信息筛选算法在保证协作信息正确率的前提下同时降低了个体间的通信量，提高通信效率；其次利用筛选出的信息对后续状态进行预测并根据预测结果制定协作策略，保证了协作的正确性；最后协作个体根据收益值对协作策略进行动态的修改，保证了算法的收敛性和健壮性，提高了智能体的决策水平。我们在 Unity 环境中对所提算法的有效性进行了验证，并将其与其他四类算法进行了比较，仿真结果表明所提算法的性能明显高于其他算法，即能够正确高效的完成协作救援任务，可以给救援决策的制定提供支持。

参考文献：

[1] Gerhard W. Multi-agent systems: a modern Approach to Distributed Artificial Intelligence [M]. Boston: MIT Press, 1999.

[2] Jiao W P. Multi-agent cooperation via reasoning about the behavior of other [J]. Computational Intelligence, 2010, 26 (1): 57-83.

[3] Semsa-Kazerooni E, Khorasani K. A game theory approach to multi-Agent team cooperation [C]// Proc. of the American Control Conference on Hyatt Regency Riverfront. 2009: 10-12.

[4] Suruz M, Bao N, Davide S. Nonuniform deployment of autonomous agents in harbor-like environments [J]. Unmanned Systems, 2014, 2 (4): 377-389.

[5] Baker C, Ramchurn G, Teacy L, *et al.* Factored Monte-Carlo tree search for coordinating UAVs in disaster response [C]// Distributed and Multi-Agent Planning. 2016.

[6] Provenzi M D O, Götz M. Activity prediction in a smart environment using Bayesian network and multi-agent system [C]// Computing Systems Engineering. 2017: 140-146.

[7] Saldana D, Assunçao R, Hsieh M A, *et al.* Cooperative prediction of time-varying boundaries with a team of robots [C]// Proc of International Symposium on Multi-Robot and Multi-Agent Systems. 2018.

[8] 吴坤，刘玮，李爽，等. 一种动态环境下多 Agent 的协作方法 [J]. 武汉工程大学学报, 2017, 39 (2): 186-192. (Wu Kun, Liu Wei, Li Shuang, *et al.* A multi-agent cooperation method in dynamic environment [J]. Journal of Wuhan Institute of Technology, 2017, 39 (2): 186-192.)

[9] 周丰，鲜明，肖顺平. 基于多智能体技术的协同信息融合系统研究 [J]. 指挥控制与仿真，2006, 28 (4): 13-16. (Zhou F, Xian M, Xiao S P. Research on co-operative information fusion system based on multi-agent technology [J]. Command Control & Simulation, 2006.)

[10] 吴志新. 基于多智能体的信息融合系统 [J]. 计算机与数字工程，2013, 41 (7): 1074-1077. (Wu Zhixin. A multi-agent based information fusion system [J]. Computer and Digital Engineering, 2013, 41 (07): 1074-1077.)

[11] 沈杰，梁志伟，刘娟，等. RCRSS 中的多警察智能体协作 [C]// 中国控制与决策会议. 2013. (Shen Jie, Liang Zhiwei, Liu Juan, *et al.* Multi-agent collaborative approaches in RCRSS [C]// Chinese Control & Decision Conference, 2013.)

[12] 彭军，刘亚，吴敏，等. 基于状态预测的多智能体动态协作算法 [J]. 系统仿真学报, 2008, 20 (20): 5511-5515. (Peng Jun, Liu Ya, Wu Min, *et al.* Dynamic Cooperation Algorithm in Multi-Agent System Based on State Prediction [J]. Journal of System Simulation, 2008, 20 (20): 5511-5515.)

[13] 王颖. 基于高斯模型的有毒气体瞬间泄漏伤害范围 [J]. 消防科学与技术, 2010, 29 (11): 1002-1004. (Wang Ying. The injury area of the toxic gas instant leaking based on Gaussian model [J]. Fire Science and Technology, 2010, 29 (11): 1002-1004.)

[14] 郑延斌，安德宇，李娜，等. 一种应用于火灾环境路径规划的蚂蚁群算法 [J]. 山西大学学报：自然科学版，2017, 40 (4): 690-701. (Zheng Yanbin, An Deyu, Li Na, *et al.* Ant Colony Algorithm for Path Planning in Fire Environment [J]. Journal of Shanxi University: Natural Science, 2017, 40 (4): 690-701.)

[15] 魏志鹏. RoboCup 救援仿真中异构多智能体协作策略 [D]. 南京：南京邮电大学，2016. (Wei Zhipeng. Heterogeneous Multi-Agent cooperation strategy in RoboCup Rescue [D]. Nanjing: Nanjing University of Posts and Telecommunications, 2016.